DOCUMENT RESUME

ED 321 362 EA 021 977

AUTHOR Collins, Angelo
TITLE Transforming the Assessment of Teachers: Notes on a Theory of Assessment for the 21st Century.
SPONS AGENCY Carnegie Corp. of New York, N.Y.
PUB DATE Apr 90
NOTE 19p.; Paper based on presentation at the Annual Meeting of the National Catholic Education Association (Toronto, Ontario, Canada, April 1990).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Elementary Secondary Education; *Evaluation Methods; *Evaluation Problems; Evaluation Research; *Portfolios (Background Materials); Professional Development; Standards; Teacher Certification; Teacher Effectiveness; *Teacher Evaluation; Teacher Qualifications

ABSTRACT

The Teacher Assessment Program (TAP) presents alternative modes of teacher evaluation to inform the deliberations of the National Board for Professional Teaching Standards (NBPTS). Three experimental methods of assessment--simulation exercise, portfolios, and simulation-based portfolios--were implemented by various groups of teachers. Practical application and observation helped to identify problems and effectiveness. Findings indicate that the methods are useful for teacher assessment and for establishing teachers as a community of professionals. (LMI)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED321362

# TRANSFORMING THE ASSESSMENT OF TEACHERS
# NOTES ON A THEORY OF ASSESSMENT FOR THE 21ST CENTURY

Angelo Collins
Stanford University

Paper Based on a Presentation at the
Annual Meeting of the
National Catholic Education Association
Toronto
April 16, 1990

Running Head: Notes on

The research reported here was supported by a grant from the Carnegie Corporation of New York. The opinions expressed are those of the author and in no way reflect the views of that organization or the National Board for Professional Teaching Standards.

For further information on the Teacher Assessment Project write: Teacher Assessment Project, CERAS 507, Stanford University, Stanford, CA 94305-3084

EA 021 977

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Angelo Collins
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

# TRANSFORMING THE ASSESSMENT OF TEACHERS: NOTES ON A THEORY OF ASSESSMENT FOR THE 21ST CENTURY*

ANGELO COLLINS
STANFORD UNIVERSITY

I'd like to thank you for the opportunity to be here with you today. As the Teacher Assessment Project at Stanford began in 1986 and will be completed in August 1990, there is something very frustrating about being asked to talk about the Project in 20 minutes. And so I appreciate the amount of time you've given me to talk about the work I do. What I'd like to do today is tell you a story that has two parallel strands--about the National Board for Professional Teaching Standards and about the Teacher Assessment Project at Stanford University. In each of the story lines you will hear some notes about a theory of assessment for the next century.

## The National Board for Professional Teaching Standards

In 1983, the National Commission on Excellence in Education in the United States issued a report entitled, A Nation at Risk: The Imperative for Educational Reform. No doubt everyone here has heard of it, many of you may have seen it, and surely some of you have read and studied this report. In 1986, the Carnegie Corporation of New York, published a response to this report entitled A Nation Prepared. This report was one of several responses to a national call for educational reform. In this report, the Carnegie Corporation called for a process to identify and acknowledge master teachers, who truly are professionals, who meet "bold, new standards of excellence in education." I am sure all of you know what a professional is, but in case you've forgotten, a professional is a person who understands the theory of teaching; a professional is one who has understanding and skill in the practice of teaching. I am not being facetious when I say professionals by definition are committed to service, and Catholic School teachers have always had a commitment to service. Lastly, other service groups that are recognized as professionals in the United States are self-regulated. Therefore, if teachers are to become recognized by the public as professionals, the professionalization of teachers needs to be by and for teachers. However, not only did Carnegie write a report with a proposal for recognition of professional teachers, they also formed a commission to do something about it and they pledged 5 million dollars to begin the design and operation of the National Board for Professional Teaching Standards. (I have told friends on the Board that they need to change the name, as NBPTS is certainly not a

memorable acronym.) Let me state loudly and clearly, that I am not a member of the National Board. I am not on the staff of the National Board. When I speak, I am speaking as a member of the Stanford University faculty and as Director of the Teacher Assessment Project, and I do not speak in the name of the National Board. But my work would not exist if the National Board was not there, and the Project and the Board are in regular communication.

In 1986, the 16 member Task Force on Education and the Economy of the Carnegie Corporation met to design a process to identify persons who would be members of the first National Board for Professional Teaching Standards. The membership of this Board is not random. There are teachers from the inner city and from rural areas; there are teachers from Alaska to Florida; there are first grade teachers and high school art specialists. (I recall David Mandel, from the Task Force, calling me one day to ask if I knew a Catholic priest who taught German on an Indian Reservation--that would fill three cells on his matrix.) The National Board came into existence in May of 1987. They are a group of 64 people from across the United States; 33 of them are classroom teachers. About the remainder of the Board, 1/2 are educational professionals such as superintendents and teacher educators and the rest come from other fields such as business and government. The National Board has worked for three years to define bold, new standards of teaching excellence and a process by which a teacher would demonstrate that he or she has met these standards. It is their intention that in 1993, the first National Board certified teacher will be recognized. National Board certification is to be voluntary. But what does Board certification imply? Maybe some of you know an accountant. There are accountants and there are Certified Public Accountants. There are genetic counselors and there are National Board Certified Genetic Counselors. Certification is different from licensure. A license merely means that the state has given you the right to perform a task and that you know enough so you won't do harm. There are licensed beauticians and morticians as well as licensed teachers. The certification would be a voluntary, national acknowledgement that not only does a teacher not do harm in a classroom, but that she does positive good.

From 1987 until July of 1989, the Board met quarterly in three committees: the standards committee, the policy committee, and the assessment committee. In January of 1989, I was privileged to attend a meeting of the National Board as a visitor. At this meeting, the Board voted on the first draft of the core propositions of what a professional teacher should know and be able to do. It was, for me, a deja vue experience, as it truly was a "board meeting." After meeting in committee, the 64 members sat at tables arranged in a large square and voted on the five core propositions. For those of you who are at least my age and are members of a religious congregation, you will recognize the

procedure as the one we used to rewrite our constitutions. The chair, Jim Kelly, read the first paragraph and asked for a vote. Then members argued words, examples, phrases. Finally, the vote was taken. The entire document was approved line-by-line.

By now you are probably curious as to what these bold new standards of teaching represented by the five core propositions, might be. Each can be stated as a sentence. Each is explained by two paragraphs, and each is supported by a two page elaboration. The five core propositions in sentence form are:

- Teacher are committed to students and their learning.
- Teachers know their subject matter and how to teach those subjects to students.
- Teachers are responsible for managing and monitoring student learning.
- Teachers think systematically about their practice and learn from experience.
- Teachers are members of learning communities.

As these propositions have been the premises on which Catholic schools have been built and developed, my hope is that in 1993, excellent teachers from Catholic schools will be among the first Board Certified Teachers.

In July, 1989, the Board published a report of its first two year's work, Toward High and Rigorous Standards for the Teaching Profession. This report includes the standards and their explication, the areas of concern that the policy committee will address, and the designation of the areas in which it will be possible to be Board certified. In February, 1990, the Board's Assessment Committee presented a draft of a Request for Proposals, to ask persons concerned about assessment to design the first assessment procedure to certify that teachers of literacy to early adolescent students have met these bold standards. It is expected that responses to the Request for Proposals will be read and evaluated by August, and that the work of designing assessment procedures to meet these standards will begin in September. The Board is creating Standards committees of teachers in each certificate area to clarify the standards in that area. Possibly some of you might be interested in joining such a committee. Last week the Board released a second Request for Proposals for researchers to examine various assessment technologies.

## Teacher Assessment Project

At this point, you are now current on the work of the National Board and we need to return to 1986 and pick up the parallel strand of the story on transforming the assessment of teachers. In 1986, Carnegie Corporation realized that, even as the Board would be struggling with defining standards for excellence in teaching,

someone would need to explore better ways to assess what teachers know and do. Thus began the Teacher Assessment Project at Stanford University. The goal of the Project has been to explore alternative modes of teacher assessment.

Our current methods of evaluating teaching and teachers are not sufficient. You may recall several years ago there was an article in The New York Times about a reporter who took the NTE and got a very high score. His article began "If I can pass this teacher test, the test has got to be a joke." The goal of the Teacher Assessment Project has been to design modes of assessment that are realistic and rigorous and demonstrate that teachers have met the bold new standards. As director of a research project, I have had one luxury that many do not have. My mandate has been to explore alternate modes of teacher assessment. And I won't, but I have the latitude to stand here and say "The things we've tried don't work."

## Assumptions

The Teacher Assessment Project (TAP) began with a number of assumptions. One is that teaching is a complex activity. As soon as we try to take the act of teaching and break it up into pieces, for a portfolio entry or for an assessment exercise, something is lost.

TAP also recognizes that there is no one right way to teach. There may be forms of teaching that are very effective that I can't imagine, because I've never seen them. Therefore, we've tried to remain open to as many varieties of teaching as possible.
TAP also believes that teaching takes place in a context. This context includes the subject matter. Teaching kindergarten is different from teaching high school physics. Teaching takes place at a particular time. The examples that I used ten years ago to teach about the immune system would certainly not be the same ones I would use if I were teaching it this year. Teaching takes place in a place; and I will pick on Kansas. Manhattan, Kansas is in the middle of the prairie. One of the most popular biology course at Manhattan High is oceanography. They are less than two miles from the Konza Prairie, a federally funded prairie restoration that goes for miles and miles and miles. There is not a single course in middle school or high school solely on prairie ecology. So teaching takes place in a place. And you're teaching persons. You know as well as I do that the 3rd period class is different from the 5th period class, and this year's class is different from last year's class. And so this contextualization of teaching drives the research we're doing at the Teacher Assessment Project.

Those two assumptions lead us to recognize that the assessment of teachers is going to require a battery of modes of assessment. No

one mode of assessment is going to be sufficient. To assess teachers we need written tests, but TAP isn't doing any research on that because other people do. Assessments also include records of direct observations. Many states already have direct classroom observations and checklists. We are not doing research on that. For assessment purposes we need records of teachers' educational activities such as transcripts. At TAP we have explored three modes of assessment: simulation exercise, portfolios and portfolios, based in simulation exercises.

Our third assumption is that the assessment of teachers needs to be done by and for teachers; therefore for every researcher on the TAP staff, there is a local teacher. And the local teachers are paid what I am paid when I work as consultant. These teachers are professional colleagues.

Our last assumption is that teachers do have both theoretical and practical knowledge. But just as expert chess players can't tell you what they do, and just as expert physicists can't tell you what they do, expert teachers can't tell you exactly what they do either. It has become so much a part of their lives that they no longer know what it is they do. And it's only through careful observation by experienced people that we, teachers and researchers together, will be able to codify the theoretical knowledge that teachers have accumulated over the years.

## History and Mathematics

With the assumptions clearly stated, let us turn to the work of the Project. The Teacher Assessment Project has five parts: mathematics, history, biology, literacy and teacher education.

In 1986, the Teacher Assessment Project began its research on new modes of teacher assessment by designing simulation exercises for teachers of mathematics and history. In the hope of avoiding some curricular controversies, the Project tried to have a very narrow focus--teaching fractions to fifth graders and teaching the American revolution to high school juniors. These two subjects were chosen because of perceived differences in the context. The elementary school teacher is seen as a content generalist, the high school teacher as a specialist. Mathematics is seen as a rigorous discipline leading to answers; history is seen as providing opportunities to discuss ideas. There is a large body of research on teaching mathematics while there was at the time relatively little research on teaching history.

The Project began by having researchers sit in classrooms of high school history teachers and fifth grade math teachers. Working with the teachers, the researchers tried to identify what we have come to call the wisdom of practice--the knowledge and skill of

experienced teachers that make their performance of teaching appear simple to non-teachers. While working with the teachers, we designed assessment center exercises, simulations of critical tasks of teaching in which teachers would have opportunities to demonstrate their expertise. During this design phase we also invited experts in assessment, in law, medicine, and architecture to visit us. Most of the exercises consisted of a task followed by an interview. For example, the high school teachers were given five essay tests--real tests done by real students--to grade and then were interviewed on how and why they made the decisions they did, how the grades on the test would influence the teaching of the next unit, and how the grading of the test might influence the teaching of the same topic next year. Other exercises in history included student questioning, using original documents, selecting a textbook. In mathematics there were exercises on checking homework, using instructional tools, and mapping concepts. One of the exercises that was designed in both mathematics and history was called "Teaching a Familiar Lesson." For this exercise, we hired students of the appropriate ages and asked teachers to teach a familiar lesson. The goal of the research was to explore how these exercises could capture the wisdom of practice. It is so easy for me to say, "We designed an interview." It took months to design the questions for the interview protocol so that they were clear and explicit and probed what a teacher knew without giving away any answers. During the summer of 1987, 20 teachers of mathematics and 20 teachers of history spent a week each at a school we had rented and completed the simulation exercises.

The TAP staff then took a year to design a rating system. Recall that the National Board at that time had not yet determined the standards. Therefore the goal of TAP was to design a scoring system that as 'data-drive'. By that I mean that rather than set up the criteria ahead of time, we would look at the data and see what the performances of the teachers brought to the surface as criteria for excellence. The result of this research on scoring is the discrete scoring system reported in "The Hundred Statements Project." (Kerdeman, 1989)

## Literacy and Biology

We had barely finished our research in 1987 when critics said "That's great. But the assessment center exercises were not really fair to teachers because they didn't have the opportunity to show what they do in their context. How are you going to capture what they do in their own schools?" And because we had the mandate to explore, we sat down and asked ourselves, "What could we do to capture the knowledge, skills and dispositions of teachers in their local situations." We began talking about portfolios.

## Images of a Portfolio

When we began our research on using portfolios for teacher assessment, we had four images of portfolios. In all instances, a portfolio is a container; it's a container of documents, and these documents provide evidence, in this case of the knowledge, skills and dispositions of a teacher. The question was, what would a teacher's portfolio look like?

One of our images was the image of the portfolio of an artist, an architect, a model. What goes into this portfolio is the artist's best work. They don't show their pen-and-pencil sketches. They know it's their best work because in some way it's been juried, formally in a show, or informally through talking to colleagues. Artists don't put everything in their portfolio. They select. And somehow through their tradition they know what it is to select, what's the right stuff. Also the artists portfolio is not static, things change in the portfolio. They change depending upon the purpose. We talked to architects about their portfolios, and they said "Watch out. There are two ways that portfolios get boring: they're boring if they've got too much stuff in them, and they're boring if they've got too little." One image of a portfolio was of selected, best work.

Another image of a portfolio is the log of a pilot. It is not a sample of the pilot's best work; it's a sample of every single thing she does. Pilots can't just hand in a record of their pretty flights. And someone else has the job of making a judgment. The pilot has no control over choosing what's in the portfolio. It's a different image of portfolios.

I'm sure you won't like this image, especially those of you who are bombarded by salespersons in your classrooms. The catalogue of a salesperson is in effect a portfolio, and it doesn't show his own work at all. What it shows is the ability to deliver someone else's work.

A final image that has great appeal is the badge of a Boy Scout or Girl Scout. Think about the badges. They demonstrate that the work has been done with a leader, in a group. It's not solo work. It's a very defined task that has been accomplished. And it's presented symbolically; and the symbol doesn't mean a whole lot to anybody except other people who are part of the group. And finally, the badges of a Scout are presented with great ceremony and great meaning.

Two other images of portfolio that have influenced our work are the tenure case of a university professor--that contains best work, work in progress and affidavits from colleagues--and the defense of a dissertation in which the student acknowledges the help of others but defends the work as his own. What is the image of a teacher's portfolio? I would hope that ultimately it would have some of the

elements of an artist, some of the elements of a catalogue, and some of the elements of the badges of a Scout. This section of the paper on images of a portfolio is based on the work of Tom Bird (Bird, 1989).

Issues in Portfolios

Size - When we started we asked ourselves questions about the size of the portfolios when we started. We really had very little idea of what a teacher's portfolio was going to look like. One person on our staff kept talking about the "boxcar approach." Can't you see all the teachers in July pulling up with their boxcars? As a matter of fact, everyone got her portfolio into a box about the size of a milk crate. I don't think that's unreasonable size.

Kinds - When we began we talked about what kinds of pieces of evidence might go into a portfolio. We thought that the teacher should have the right to choose. Some of the pieces should definitely be elective. What do you think is the best evidence of your own practice of teaching? On the other hand, some of the evidence should be prescribed.

Assistance - Because we believe that teaching is by and for teachers, we wanted the pieces of a portfolio to be guided. Therefore, everyone had an advisor in the portfolio development process. And in some instances the portfolio contained attestations that this was the teacher's own work. People ask me "How do you know the work in the portfolio is really the teacher's work? How do you know they didn't just write it down to make it look pretty? What about fraud?" And the teacher part of me gets very sad that teachers don't trust one another. My brother is a physician. When he walked in with his caseload to get his national board certification in medicine, nobody said to him "Dr. Collins, are those really your cases?" And so when a teacher walked in with 5 samples of tests she had given this year, I didn't say "Gee, I wonder if she really gave them." On the other hand, in order to respond to the national concern, teachers were interviewed about things in their portfolios, so if the materials weren't their own, they were going to have a very difficult time talking about them.

Purpose - We talked about what purposes a portfolio might have. One vision of portfolio is for summative evaluation, something that a teacher would have at the end of three-to-five years of teaching. Beginning about the junior or senior year of college, going through a fifth year program, a formal internship, and three to five years of teaching they would produce a portfolio. I know that portfolios are formative. I have worked with 45 teachers for a year, and I have listened to them talk about what they have learned about themselves as teachers, about what they have learned about the art of teaching, about what next year is going to look like because they've done this portfolio. I have no doubt about

portfolio's formative value. I know I have a dream about the celebratory value of portfolios--that someday the presentation of the teachers portfolio will be a sign of pride. This year, a teacher went to another district looking for a job and brought her portfolio along. The principal and superintendent who were interviewing her said "We've never seen anything like this before." I'm sure portfolios have an organizational value.

Types - What goes into a portfolio? Four things.

> Artifacts. Stuff teachers really produce,--tests, worksheets, notes to parents, scribbled notes about dangerous things in a biology laboratory experience.
> Reproductions. Videotapes, audiotapes, photographs of bulletin boards and blackboards--the materials teachers usually don't capture.
> Reflections. Reflections are the things that teachers wouldn't ordinarily do, such as written explanations and reasons for decisions. But teachers do them in portfolios because they're asked to do them, and reflections end up being valuable to teachers in thinking about their own practice.

One of the teachers that I worked with last year said that one of the neatest things about portfolios was this was the first time anybody had ever looked at what he did. And it was just so important to him to realize that somebody else thought this stuff was important. He said "I've been producing this stuff for 26 years, and no one has ever asked to see a sample of an overhead transparency. No one has oohed and aahed at this biology bulletin board that students contribute to that I have perfected over the years. If nothing else came out of the portfolio but to have someone else treat my work seriously, it was worth it."

Summary - Those were our ideas when we started to design a teacher's portfolio. We summarize them like this. We wanted a portfolio that a teacher could do during one academic year. (That's a mistake. A year is too short.) We wanted to make sure that the portfolio process we designed required that all teachers work with other experienced teachers, so the portfolio would be an opportunity to form collegial relations. We wanted to encourage teachers with various contexts and various styles. One of the things that we had found out in the assessment center exercise was that teachers were very good at telling us what they would do or what they might do. In the portfolio we wanted a sample of what they actually did do. And we wanted to design experiences that would enhance teachers' growth.

Design

My colleague Dr. Linda Vavrus, now at the University of Nebraska at Lincoln, sat down in the summer of 1988 with six Stanford

researchers and six people who have experience in teaching 3rd and 4th grade and asked "How are we going to capture the knowledge and skills of 4th grade literacy teachers? What might a 4th grade literacy teacher's portfolio look like?" They chose to focus on literacy--the integration of reading, writing and communication--at third and fourth grade because that is where the emphasis shifts from learning to read to reading to learn. They also chose to focus on literature-based instruction. They finally decided that the way to cut up teaching 4th grade reading is in three parts. They called their three parts scenarios, because they represented things they saw happening in classrooms. They called their parts "Teaching an Integrated Language Lesson," and since I'm not a teacher of literacy I have to take it on faith that that's very important; "Assessing Student Knowledge," and I have heard lots of talk about assessing student knowledge at the beginning of the year to design a reading program, during the year to keep the reading program on track, at the end of the year to make recommendations to the next teacher. The last portfolio entry they designed was "Creating a Literate Environment." That one caused them trouble, because somehow or other they knew what it was but not exactly how to capture it. It had to do with where the books were, and how tattered and torn they were, and when the kids could check them out, and if there was a reading center. But it also had to do with creating the attitude that books were important. The question was "What kind of evidence would you collect to demonstrate that a teacher had created a literate environment?" This is one of the places they used videotape; they used photographs; they talked to students; they had samples of student work.

As the research was conducted in parallel, I will concentrate on the work in biology for the remainder of the paper. While Dr. Vavrus was sitting with her teachers, I was sitting across the hall with the high school biology teachers, and they said "No problem. We can divide high school biology teaching into four critical tasks: Instruction, Planning and Preparation, Evaluation and Reflection, and that other thing." We began called "that other thing" "Exchange." I think last week somebody said it should be changed, that word should be "Involvement." Later I will tell you what I mean by "Exchange," or "Professional Development."

We had no trouble discussing planning. We generated this long list of planning events--planning units, planning for the year, planning for the week. We finally decided that for our experiment in portfolios we would ask the teachers to submit evidence of how they do a unit plan. So we prescribed that there had to be a unit plan, but teachers elected what unit they wanted to document. Our definition of "unit" was "a sequence of instruction bounded by a natural event, such as a week, a test or an experiment or a topic."

Designing the Instruction Portfolio Entry was fun. Again, we generated a list of possible instructional activities. High school science teachers do laboratories. They've got to use something besides the textbook for instruction. What about a lecture? What about a cooperative small learning group? What about--and our list again got long. And finally we decided that teachers could submit evidence of how they do a laboratory lesson or how they use alternative materials. They were these two choices of instructional activities from a list of 17. They represented two very different but very important approaches to instruction.

For Reflection and Evaluation we began thinking about reflection about and evaluation of students. The first item on the list was: How do you design a good test? How many different kinds of evaluation methods do you use for your students, and their knowledge and skills? What about evaluation and reflection on lessons? What about evaluation and reflection on my teaching? Again we had our list, and finally decided that what we were going to put in the portfolio was the teacher's log of the variety of methods used to evaluation student knowledge.

Then the research team got very uncomfortable, because there were still two things missing in our portfolio design. Earlier we had consciously decided to limit the number of entries in the portfolio because we knew that teachers could only do so many in the given amount of time. But there were two things that we were unhappy about. One was that we had not captured the relationship between a teacher and her students. When I went around and talked with teachers and asked them what they thought was the most important characteristic of a successful high school biology teacher, I got one of three answers: enthusiasm, creativity, or relationship with students. When I asked what teachers would want to do to demonstrate their knowledge and skills if they were going to design the portfolio, they always wanted something to do with student-teacher relationship. And I will stand here and say, with shame, that we have not in four years of research come up with a way of evaluating how teachers form relationships with students. I think some of our assessments get at it through the back door, but we haven't come up with some mode of assessment that focuses on relationship. But it is not because we think it is unimportant, but because we have tried and failed.

And the other important aspect of teaching that was missing completely in the portfolio was what happens outside the classroom. The arguments went something like this: A person could be very good in their classroom, but if they don't participate in the community, whether community is defined as school or local or professional there's something missing in what they do. And as soon as we all agreed to that, someone would say "But I don't want to certify somebody who's running all over the state telling everyone how wonderful he is, but he does a rotten job in his

classroom." This discussion went on for almost two days. And without a name for it, and without being exactly sure how to define it, we stuck it into the portfolio and we called it Exchange.

We didn't do any better job after defining the area. We had been comfortable listing 17 different kinds of instruction and choosing two. But what we did when designing Exchange was list 27 different kinds of exchange and tried to design a portfolio entry that could contain them all. We violated our own principle; we went for the boxcar approach.

Ultimately a high school biology teacher's portfolio looked like this. The teacher submitted background information. As we believe in the importance of context and we wanted to know whether they were teaching in Menlo-Atherton High School or in downtown San Francisco. We wanted to know whether they had 38 students in their classroom or ten. Teachers submitted entries on unit planning, teaching a lesson using alternate materials or a laboratory lesson, assessing students' knowledge and skills, and exchange.

Doing Portfolios

In September, 1988, we met with 20 high school biology teachers (and 24 teachers of literacy at 4th grade level) who were going to develop portfolios. For financial reasons we concentrated the sample of teachers in the San Francisco Bay Area. We had teachers in large schools and small schools, we had teachers in districts that bragged that 90% of their students go on to college, we had teachers in schools where only 5% of the students go on to college. We had elementary teachers in rural districts. We had a few teachers from Ohio, Arizona and Alaska. The youngest teacher we had was an intern; the most experienced teacher had 30 years of experience. When we met we presented them with a notebook called the Portfolio Development Handbook. In this book we gave them an overview of the Project and explained that our goals were to explore a new mode of teacher assessment. In the handbook we set out images of portfolios as I did earlier in this paper and types of evidence. We defined each of the portfolio entries and included what was required and some suggestions for evidence. We defined words that we had assigned special meaning to, such as unit and laboratory. We listed the reasons why each entry had been designed and suggested some of the criteria we would use in evaluating the entries. We assigned each teacher an advisor, and sent them on their way. With all my heart I need to praise, thank, and congratulate the teachers who participated in the Project and developed portfolios. None of us had any idea what we were getting into when we began. They worked very hard in a constant state of ambiguity.

Videotapes As Evidence

Now that I have seen 44 portfolios, if you would ask me what's the one necessary bit of evidence in a teacher's portfolio, I would reply "a videotape." And as soon as I said that, I would say the videotape is necessary but not sufficient. The videotape by itself is almost meaningless, unless you know what happened yesterday, what's going to happen tomorrow, and where the taped lesson fits into some grander scheme of teaching.

When we began, the idea of using videotapes was very controversial. Issues include:

-access [if a teacher in this school district doesn't have any equipment, how will he produce a videotape]
- expertise [this school has a course in videography, and there are seniors who have been doing football games and basketball games, and this teacher's videotapes are going to be really good]
-influence on students [what's going to happen to the students when you bring in this video equipment and the students are going to behave differently because there's a videotape in the room] and
- point of view [person doing the videotape automatically decides what's going to be taped].

We had our staff do all the videotapes. And the first laboratory we taped, the person doing the taping came back to the office and said "I don't know what to do. Should I take this group of four students and have the tape on them the whole time? Or should I follow the teacher around and see what's going on? Or should I move from group to group? How do I capture what goes on in the classroom on a videotape?" And despite all the problems the videotapes are an excellent piece of evidence about teachers knowledge and skills.

If you're using videotapes for the assessment of teaching and learning, one of the things that is extremely valuable but costs a lot of time, is to have the teacher sit down ahead of time with the person videotaping, and tell them what they want to capture on tape. For example, a fifth-year teacher sat down with his advisor to write a story board of what the videographer was supposed to film in the next day's lab. The advisor said "What's important about this lab?" And they spent three hours in discussion, at which point they decided there was nothing worth capturing on the videotape; as a matter of fact, there was nothing worth doing in the whole lab. Doing the videotape provided an opportunity for two teachers to think about and discuss their practice.

## A Second Assessment Center

Teachers spent from September to June completing their five portfolio entries. The range of time it took them was from 10 hours to 96 hours for each entry. Each entry ranges in size from three pages to three 3-ring binders. As the teachers worked at completing their portfolios, the staff of the Teacher Assessment Project moved on to another phase, an assessment center for teachers of biology and literacy who were developing portfolios. One reason for this second assessment center was to try to adapt some of the exercises that had been developed for mathematics and history to new domains--biology and literacy. Another goal was to determine if we could design a new mode of assessment--portfolio-based simulation exercises--that would take advantage of the best characteristics of both simulation exercises and portfolios. That is, a new mode of assessment that would capture both the standardization of simulation exercises and the context of portfolios. The easiest way to describe portfolio-based exercises is to give an example. A teacher had given us evidence of a laboratory biology lesson in her portfolio. When she came to the assessment center we interviewed her, saying "We'd like you to recall your laboratory lesson, what was it about? Imagine that halfway through the lesson you notice that the four students at Table A weren't paying attention. They weren't doing the laboratory work. Can you tell me why you think that might happen, what are you going to do about it, and why is that the appropriate thing to do?" When the teacher completed the interview process, the interviewer asked, "Now give us another interpretation of what was going on at Table A, why you think that happened, what you're going to do about it, and why that is appropriate." Then we asked another series of questions about the students not understanding what they were doing. Finally, we asked a third set of questions about time management--it's ten minutes before the period ends and you realize nobody's going to finish. So each teacher got the same standardized questions, but they were able to situate the response in the context of her own classroom laboratory lesson.

Another goal of the assessment center held in 1989 was to employ a holistic, real-time rating system based on the five core propositions from the National Board. In less than a week, the Teacher Assessment Project rated four portfolio entries for every biology teacher, tailored three portfolio-based simulation exercises for each of the teachers, administered six simulation exercises to each teacher, rated all the exercises, and formally discussed the quality of each teacher's performance.

## Problems

Portfolio is quickly becoming a buzzword. This frightens me because I know portfolios still have problems. What are the problems with teacher portfolios? One is history. Teachers have no history of producing materials they will share with colleagues. Teachers produce materials and share them with students, and occasionally share them with parents. But normally what teachers do stays in some little file cabinet in the back of the room. So they don't have a tradition of sharing their work. Therefore, there is no collegial definition of what is acceptable. We, teachers and researchers together, need to create a tradition of what to select to put into a portfolio.

Another major tension about portfolios is defining the critical tasks that the portfolio will build around. Some teachers wanted to do all the biology portfolio entries on one topic. However, if all the entries were focused on photosynthesis no one would be sure how well the teacher taught ecology.

The third problem is the problem of display. The staff has begun to call this lamination. The question is: Is that portfolio in a manila envelope considerably poorer than the one that came in a red plastic binder? The one with the black and white photographs, is it poorer than the one that had a videotape tour of the classroom? The answer is no. The raters, again themselves classroom teachers, had no trouble separating substance from show. But the tendency to laminate is very strong. And that makes sense because we don't have a tradition of what's good and what isn't. We have no tradition of showing other teachers in our own schools, in our own districts, in our own towns, what it is we do. And since we're not sure how really good it is, at least it'll look pretty.

We have not come up with a method of evaluating teachers through portfolios that doesn't require people to be articulate. You can't do a good job on portfolios simulation exercises unless you can talk and write well. The problem gets even more complex when you start imposing cultural norms on conducting interviews. In some cultures, like mine, you want to show off and say everything you possibly can. In other cultures you answer the question, and to say more than what you are asked is an insult, because then you're assuming that the questioner isn't smart enough to infer all that you didn't say.

Another problem that we haven't been able to get to go away is the tension of teaching as it is and teaching as it might be. I know the portfolios we designed were skewed toward a vision of what teaching might be.

The fifth part of our project is called the Teacher Education Consortium. One of our concerns is that the assessments contain subtle forms of bias. We are also concerned about how Teacher Education might have to change if performance-based assessments were to become the norm. Therefore, the University of Alaska, Florida A&M University, Pan American University, City College of New York and a consortium consisting of the Dayton public schools and Wright State University, all schools that have a history of educating minority teachers, have come to Stanford several times in the last three years, and spent time with us looking at our portfolio handbooks, listening to us talk, looking at our samples. They have gone back and tried to decide what a portfolio would look like for teacher education students in their schools.

## Conclusion

At this point, the threads of the story come together. For four years, the Teacher Assessment Project at Stanford has been exploring alternate modes of teacher assessment with the intention of informing the deliberations of the National Board for Professional Teaching Standards. The National Board has been determining bold new standards for excellence in teaching. Hopefully, some of you will and be among the first teachers to be nationally certified. However, everyone of you is in a position to make some judgments about the assessment of excellent teaching. While portfolios and simulation exercises hold great promise for assessing teachers, they have an even greater value in helping teachers form themselves as a community of professionals. The value of these modes of assessment is both in the design and in the performance. To meet with a group of teachers to identify the critical tasks of teaching, to determine criteria for excellence, is to open the door to serious discussions about teaching.

Ever since I began teaching, someone has said to me, this is an exciting time to be in education. Whether you want to become Board certified, or whether you want to become better, the work on alternative modes for assessment for bold new standards is truly an exciting event.

**END**

## References


Bird, T. (1989) The schoolteacher's portfolio; An Essay on possibilities. In J. Millman and L. Darling-Hammond (Eds.), Handbook of teacher evaluation: Elementary and secondary school personnel, Second Edition.

Carnegie Forum on Education and the Economy, (1986). A Nation prepared: Teachers for the 21st century. Carnegie Corporation of New York.

Kerdeman, D. (1989) The 100 statements project: A study of the dynamics of teacher assessment. Technical Report, Teacher Assessment Project, Stanford University.

National Board for Professional Teaching Standards, (1989). Toward high and rigorous standards for the teaching profession. Washington, D.C.

National Commission on Excellence in Education, (1983). A Nation at risk: The imperative for educational reform. United States Government Printing Office, Washington, D.C.